

**Kelompok 01**

Cancer Risk Prediction



Welcome to the Cancer Risk Prediction App, designed to assess the likelihood of cancer based on various health metrics.

[Explore Now](#)



Anggota Kelompok

068

Muhammad
Ardiansyah Asrifah

065

Izzata Clarissa
Salsabila

066

Zabryna Andiny

076

Elva Aprili Timang



Data Wrangling

01

Preprocessing Data

	Age	Gender	BMI	Smoking	GeneticRisk	PhysicalActivity	AlcoholIntake	CancerHistory	Diagnosis
0	58	1	16.085313	0	1	8.146251	4.148219	1	1
1	71	0	30.828784	0	1	9.361630	3.519683	0	0
2	48	1	38.785084	0	2	5.135179	4.728368	0	1
3	34	0	30.040296	0	0	9.502792	2.044636	0	0
4	62	1	35.479721	0	0	5.356890	3.309849	0	1

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1500 entries, 0 to 1499
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Age              1500 non-null   int64  
 1   Gender            1500 non-null   int64  
 2   BMI               1500 non-null   float64 
 3   Smoking           1500 non-null   int64  
 4   GeneticRisk       1500 non-null   int64  
 5   PhysicalActivity  1500 non-null   float64 
 6   AlcoholIntake     1500 non-null   float64 
 7   CancerHistory     1500 non-null   int64  
 8   Diagnosis          1500 non-null   int64  
dtypes: float64(3), int64(6)
memory usage: 105.6 KB
```

	Age	Gender	BMI	Smoking	GeneticRisk	PhysicalActivity	AlcoholIntake	CancerHistory	Diagnosis
count	1500.000000	1500.000000	1500.000000	1500.000000	1500.000000	1500.000000	1500.000000	1500.000000	1500.000000
mean	50.320000	0.490667	27.513321	0.269333	0.508667	4.897929	2.417987	0.144000	0.371333
std	17.640968	0.500080	7.230012	0.443761	0.678895	2.866162	1.419318	0.351207	0.483322
min	20.000000	0.000000	15.000291	0.000000	0.000000	0.002410	0.001215	0.000000	0.000000
25%	35.000000	0.000000	21.483134	0.000000	0.000000	2.434609	1.210598	0.000000	0.000000
50%	51.000000	0.000000	27.598494	0.000000	0.000000	4.834316	2.382971	0.000000	0.000000
75%	66.000000	1.000000	33.850837	1.000000	1.000000	7.409896	3.585624	0.000000	1.000000
max	80.000000	1.000000	39.958688	1.000000	2.000000	9.994607	4.987115	1.000000	1.000000

03



Data Cleaning

02

Preprocessing Data

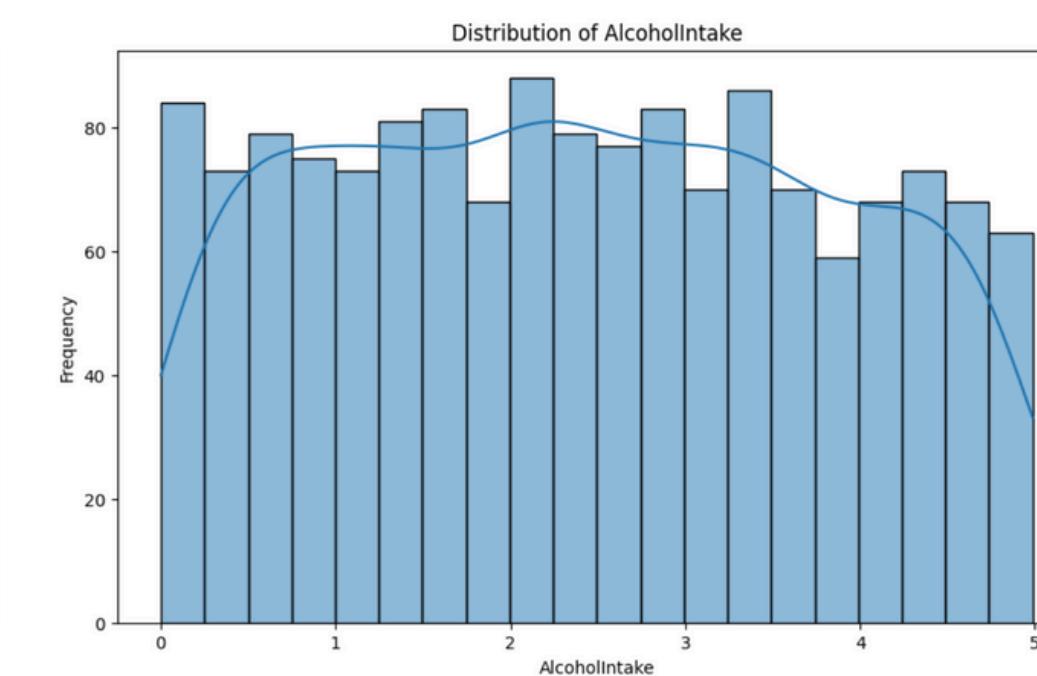
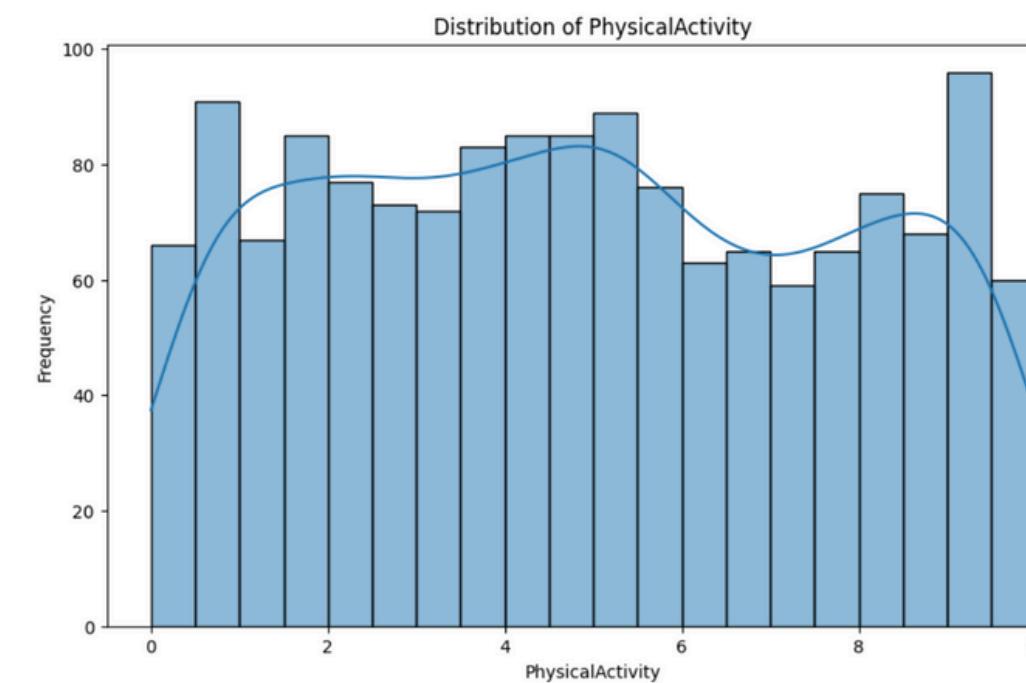
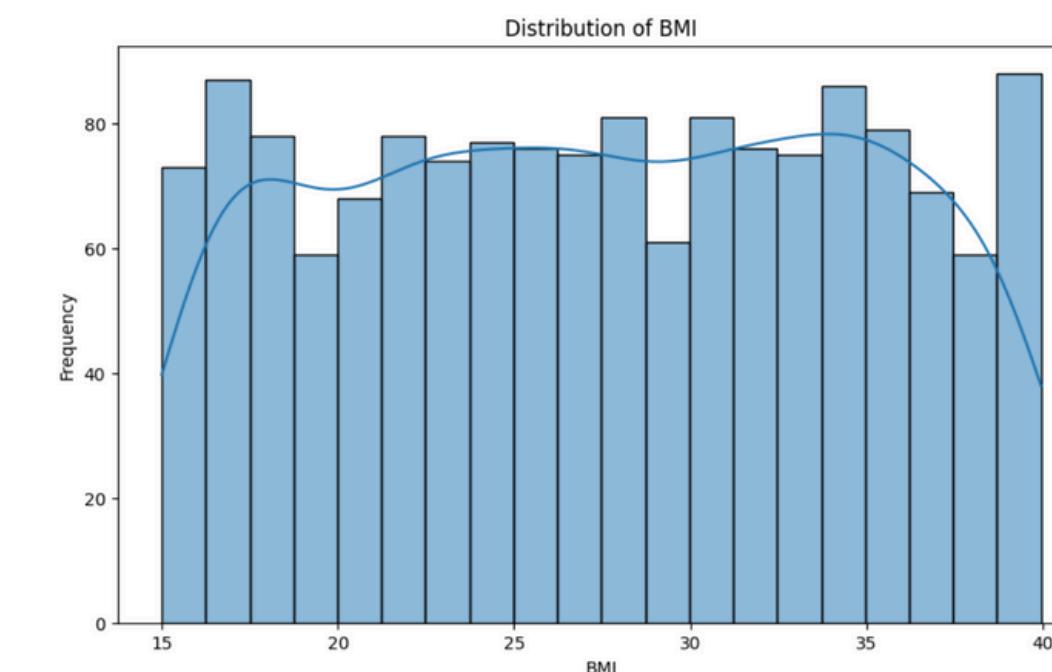
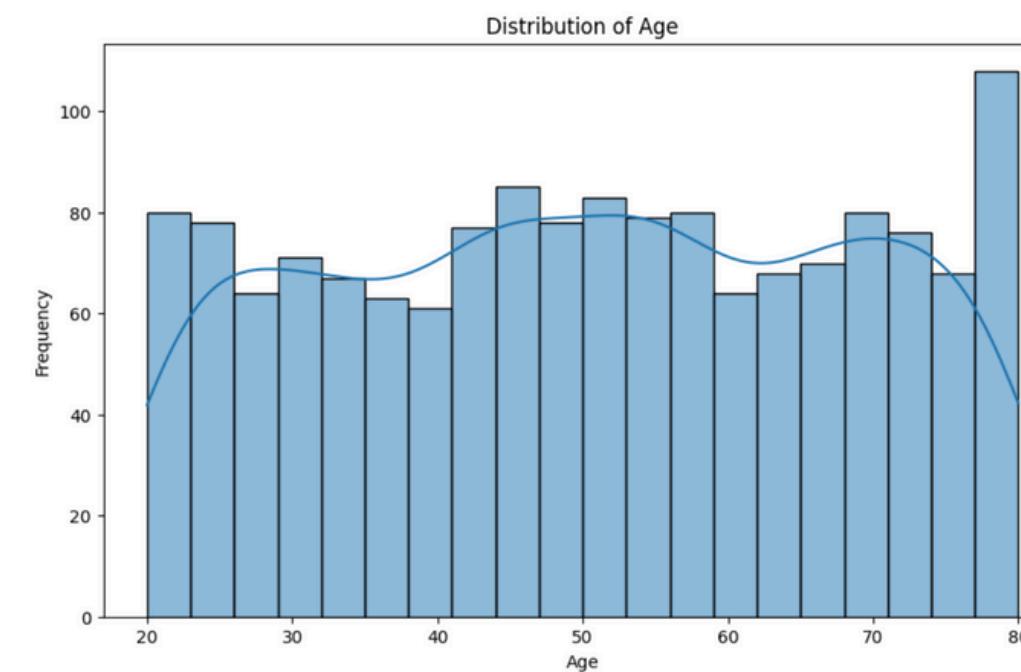
```
Age          0
Gender       0
BMI          0
Smoking      0
GeneticRisk   0
PhysicalActivity 0
AlcoholIntake 0
CancerHistory 0
Diagnosis     0
dtype: int64
```



Data Exploration

03

Preprocessing Data



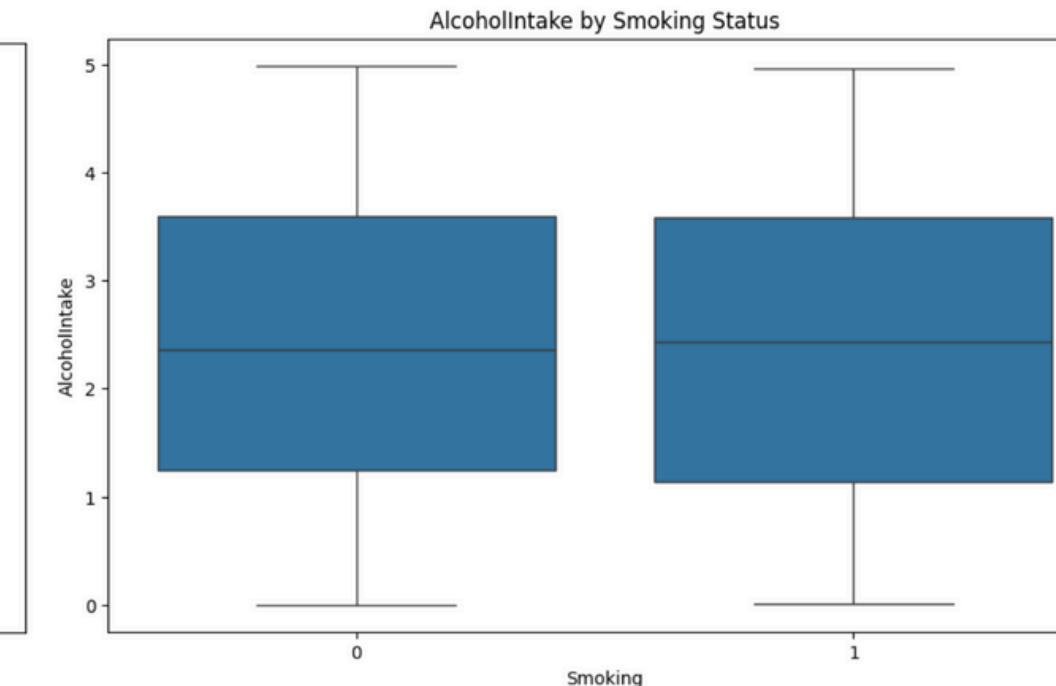
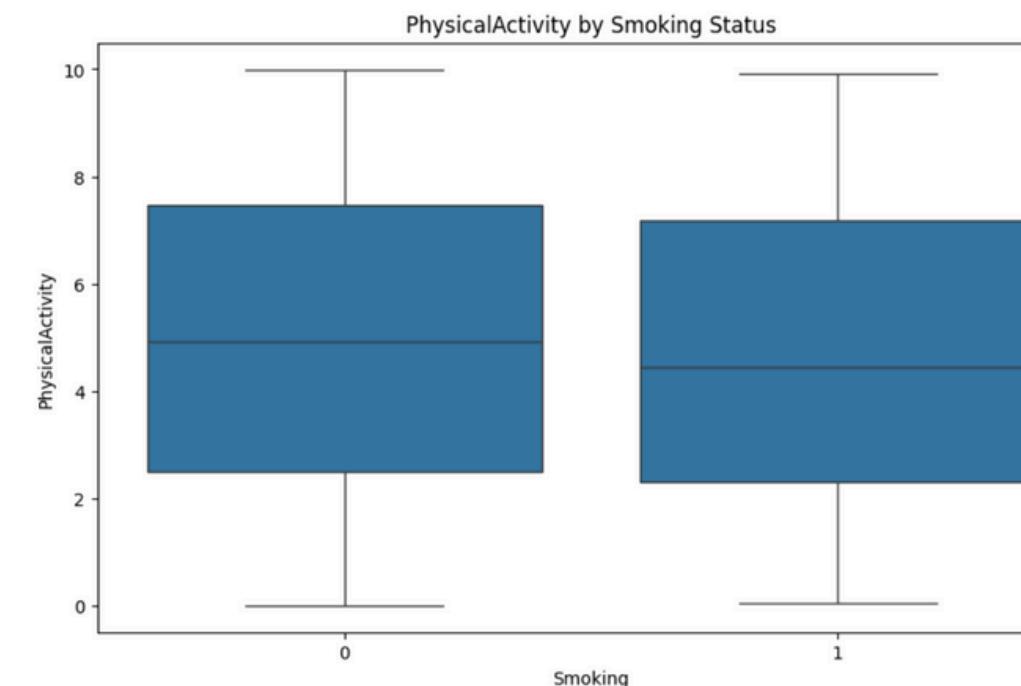
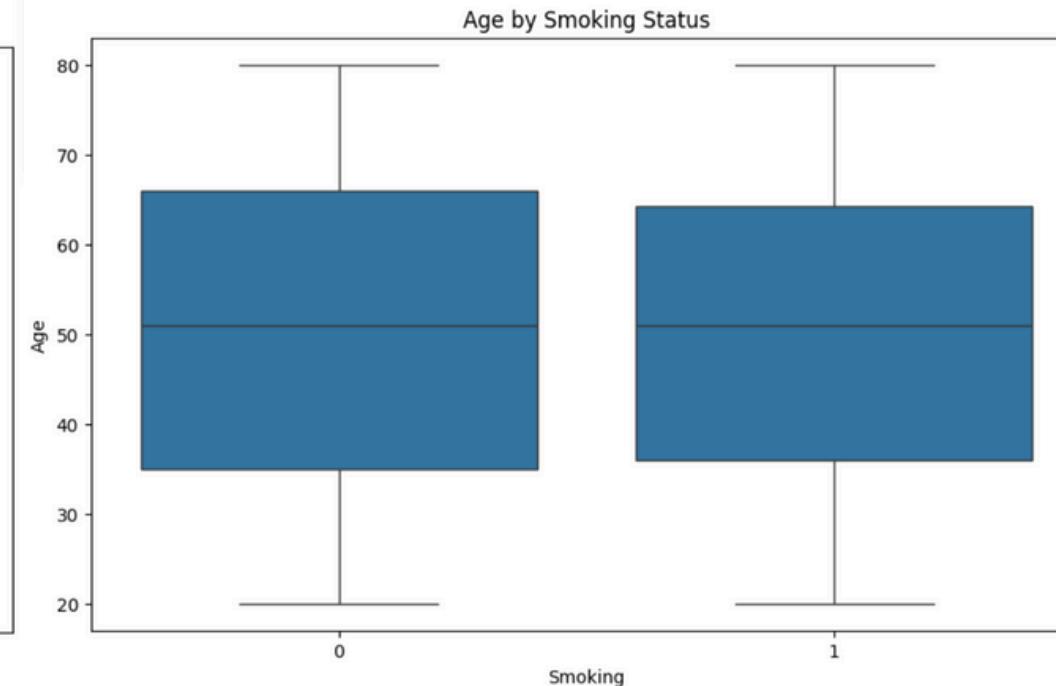
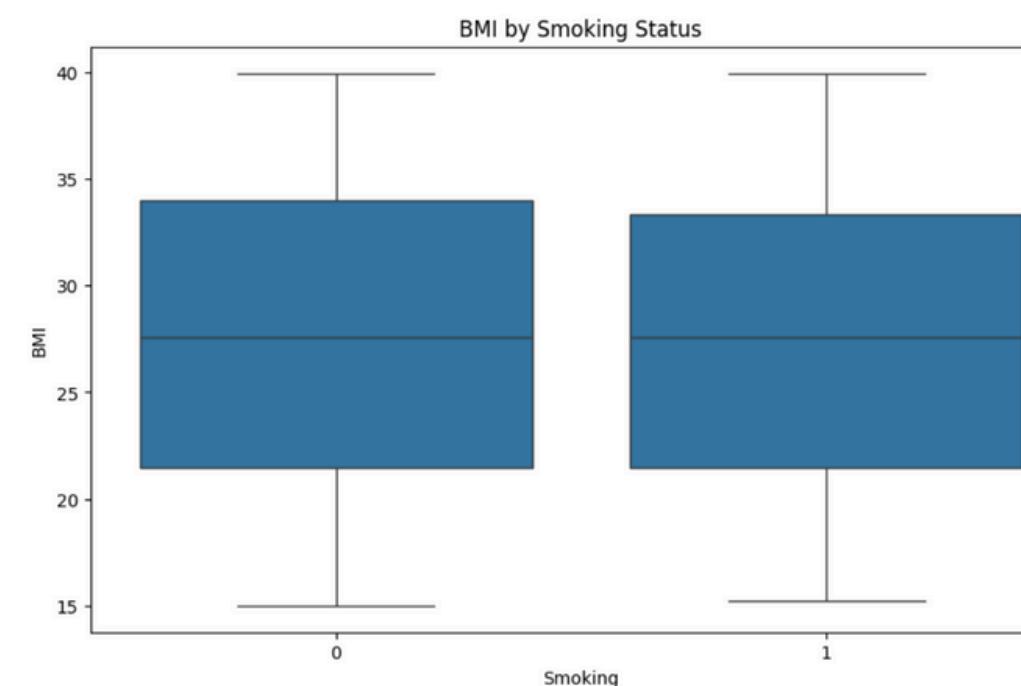
05



Data Exploration

03

Preprocessing Data

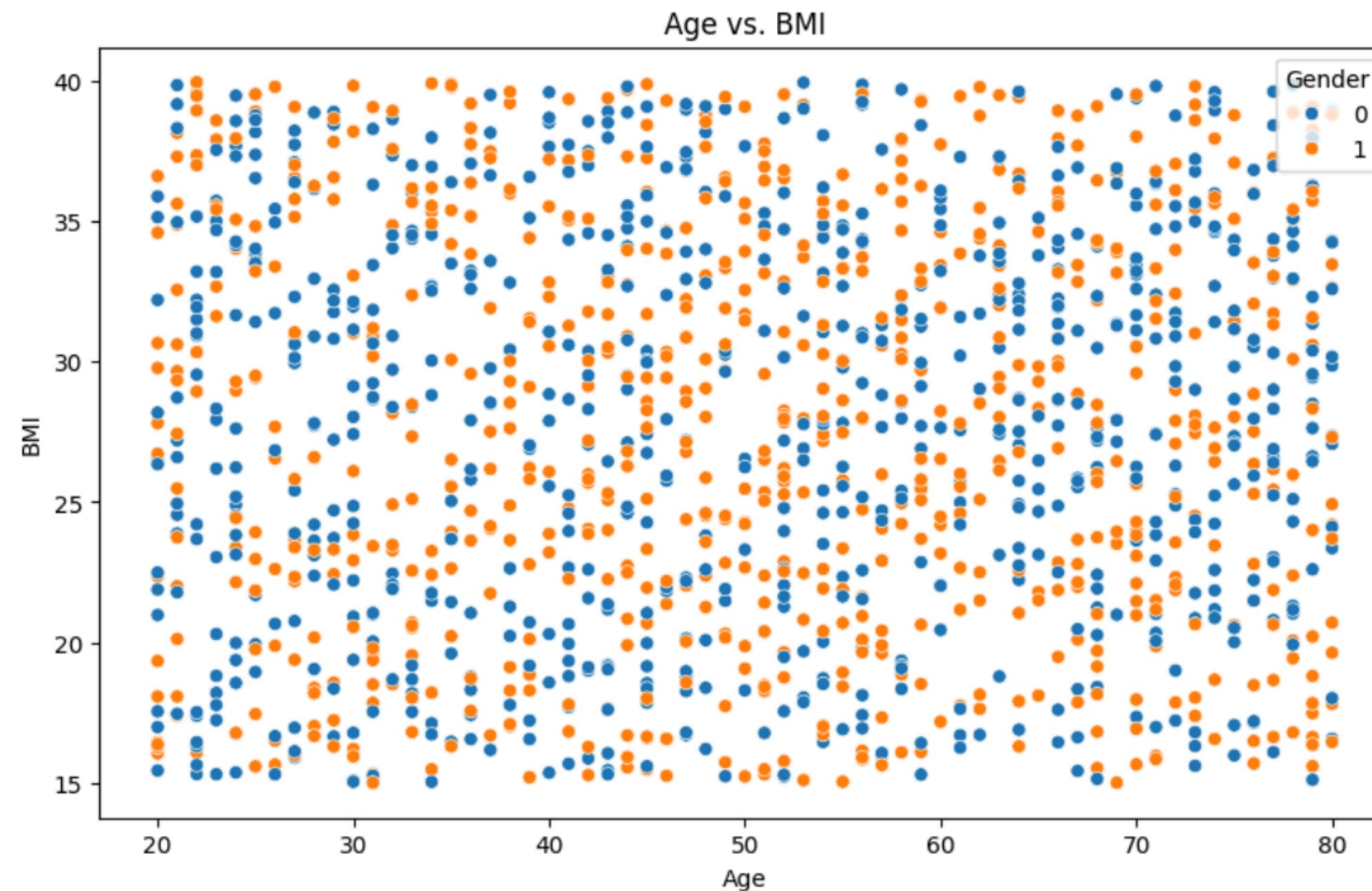




Data Exploration

03

Preprocessing Data

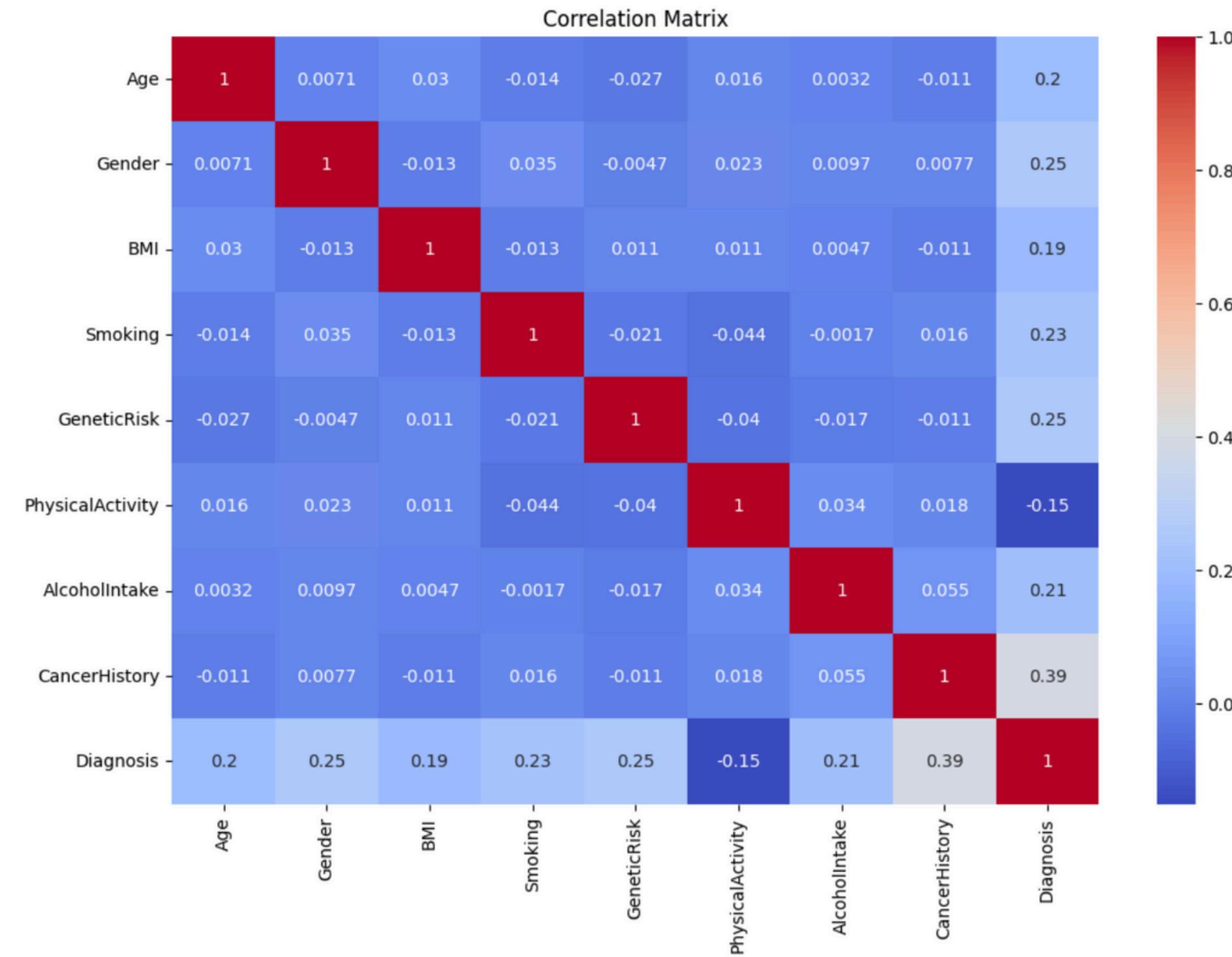




Data Exploration

03

Preprocessing Data





Structure Code



- ┌ └ assets
- └ cancer.jpg
- ┌ └ datasets
- | └ cancer_risk_data.csv
- ┌ └ models
- ┌ └ random_forest_model.pkl
- ┌ └ scaler.pkl
- ┌ └ src
- | └ train_model.py
- ┌ └ utils
- | └ app.py
- | └ bmi_calculating.py
- | └ evaluatingmodels.py
- | └ preprocessing datasets.ipynb
- ┌ └ app.py
- └ README.md
- └ user_feedback.txt



Predict

Halaman ini biasanya menawarkan navigasi yang intuitif ke berbagai bagian aplikasi dan ada penambahan fitur terkait menghitung BMI



Visualization

area yang didedikasikan untuk menampilkan grafik, plot, tabel, atau visualisasi lainnya yang membantu pengguna memahami dan menganalisis data



Service

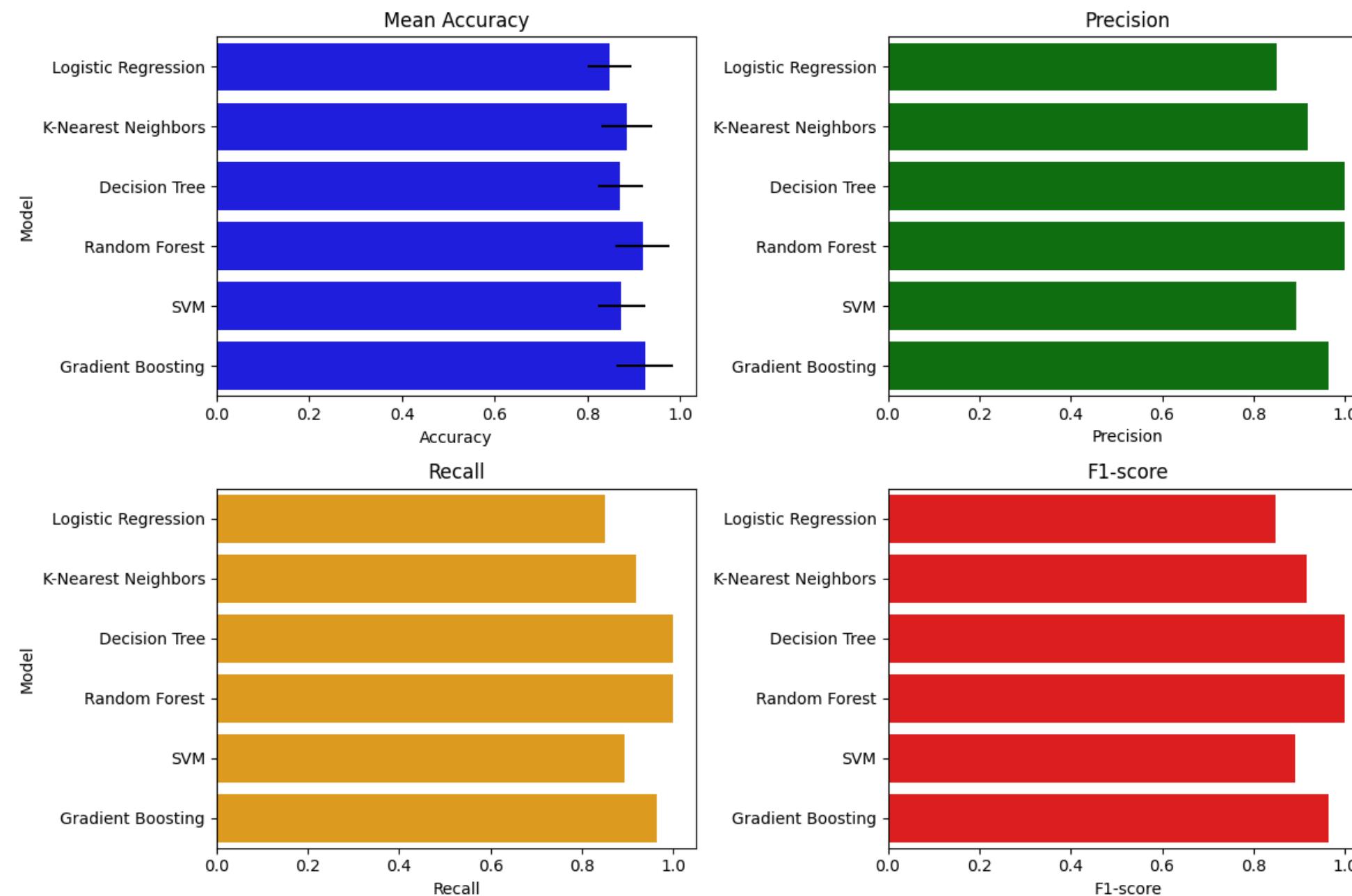
tempat di mana pengguna dapat memberikan masukan, bertanya, atau mengakses materi pendidikan yang relevan.



Fitur Fitur



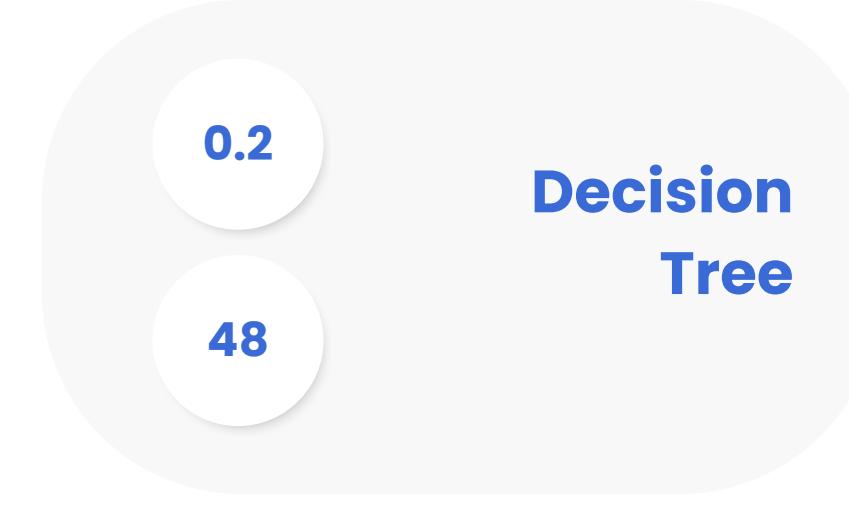
Perbandingan Model





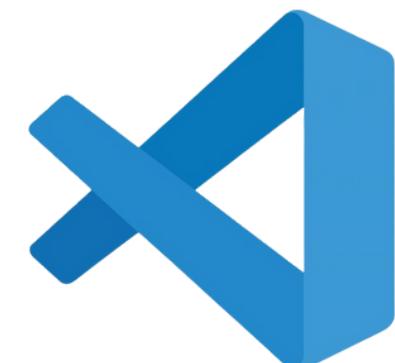
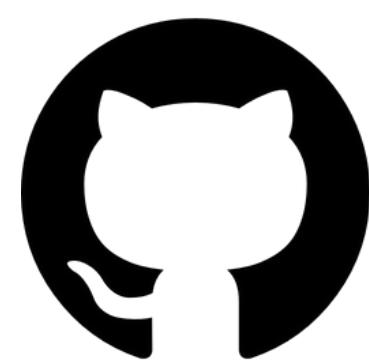
Running Time (seconds)
Model Size (bytes)

Running Time & Model Size **Perbandingan Model**





What We Use ?





More About Code ?

[Scan Our Github](#)



SCAN HERE



More About Report ?

Scan Our Report Docs



SCAN HERE



More About Datasets ?

Scan The Kaggle
Datasets



SCAN HERE



Demo Application



Thank You

